Enzyme Engineering Enables Inversion of Substrate Stereopreference of the Halogenase WelO5*

Voss, Moritz; Hüppi, Sean; Schaub, Daniela; Hayashi, Takahiro; Ligibel, Mathieu; Sager, Emine; Schroer, Kirsten; Snajdrova, Radka; Buller, Rebecca

DOI
10.1002/cctc.202201115

Publication date
2022

Document Version
Final published version

Published in
ChemCatChem

ChemCatChem

Research Article
doi.org/10.1002/cctc.202201115

Chemistry Europe
European Chemical Societies Publishing

www.chemcatchem.org

VIP *Very Important Paper*

Enzyme Engineering Enables Inversion of Substrate Stereopreference of the Halogenase WelO5*

Moritz Voss,[a] Sean Hüppi,[a, b] Daniela Schaub,[a] Takahiro Hayashi,[a, c] Mathieu Ligibel,[d] Emine Sager,[d] Kirsten Schroer,[d] Radka Snajdrova,[d] and Rebecca Buller*[a]

Enzymatic late-stage diversification of small molecules has the potential to rapidly generate diversity in compound libraries dedicated to drug discovery. In this context, freestanding Fe(II)/α-ketoglutarate-dependent halogenases have raised particular interest as this enzyme family allows the otherwise difficult regio- and stereoselective halogenation of unactivated C(sp^3)–H bonds. Here, we report the development of two engineered variants of the halogenase WelO5* for the racemic resolution of a mixture of stereoisomers generated in the synthesis of a bioactive martinelline-derived fragment. By screening a 3-site combinatorial variant library, we could identify two variants exhibiting exquisite substrate selectivity towards the desired enantiomers. Strikingly, the inversion of substrate stereopreference between the halogenase variants was achieved by varying only three residues in the active site. Protein crystallization and subsequent structure elucidation of the wildtype enzyme and a WelO5* variant shed light on the factors governing substrate acceptance and selectivity.

Introduction

Halogen incorporation alters the biological activity and pharmacokinetic profile of a molecule, making the halogenation reaction an important modification strategy in the field of medicinal chemistry.[1] In particular, late-stage halogenation of bioactive compounds is of high interest as it allows to accelerate structure-activity relationship (SAR) studies by investigating the biological effect of halogenation on existing drugs.[2] As the field of biocatalysis is maturing, new enzymatic synthesis opportunities enable the enzymatic regio- and stereoselective halogenation of unactivated carbons,[3] a chemical transformation which is challenging to perform with modern chemical synthesis approaches. While the halogenation of electron-rich carbon centers can be carried out by flavin-dependent halogenases,[1,4] Hillwig & Liu paved the way for enzymatic halogenation of aliphatic substrates by characterizing halogenase WelO5 from *H. welwitschii* UTEX B1830, the first Fe(II)/α-ketoglutarate-dependent halogenase (αKGH) found to act on freestanding substrates.[5] Up to this point, the enzymatic halogenation of unactivated C(sp^3)–H bonds was described for carrier-protein-dependent αKGHs only, which were in general of little interest for industrially relevant biocatalysis applications. With additional members of freestanding αKGHs discovered over the past years, the available toolbox for aliphatic halogenation has been steadily growing and today enables the halogenation of selected indole alkaloids,[5–6] amino acids,[7] and nucleotides.[8] Notably, protein-engineering studies of the WelO5 homologs WelO5*[9] and *Wi*-WelO15,[6c] successfully tackled the acceptance of non-native substrates. The first study, reported by Hayashi *et al.*,[9a] described the engineering of WelO5* for the selective halogenation of a martinelline-derived fragment. Here, the WelO5* halogenase from *Hapalosiphon welwitschii* IC-52-3,[6a] which showed initial chlorination activity towards the martinelline-derived fragment, was successfully evolved by targeting rationally chosen active site residues. The resulting double mutants were active and selective towards the target compound, forming two distinct chlorination products. Remarkably, one of the two variants, namely WelO5* CB2 (V81R/I161S), abolished the formation of hydroxylated side-products. In a similar structure-guided directed evolution study, Duewel *et al.* engineered the *Wi*-WelO15 halogenase towards the chlorination of non-natural hapalindoles,[6c] whereas the most recent example by Büchler *et al.* highlights the use of machine learning-driven directed evolution to create enzyme variants capable of halogenating the bulky macrolide soraphen A.[9b] In all studies, the engineered halogenases enabled the preparative synthesis of non-native chlorination

[a] *Dr. M. Voss, S. Hüppi, D. Schaub, Dr. T. Hayashi, Prof. R. Buller*
Competence Center for Biocatalysis
Institute for Chemistry and Biotechnology
Zurich University of Applied Sciences
8820 Wädenswil (Switzerland)
E-mail: rebecca.buller@zhaw.ch
[b] *S. Hüppi*
Department of Biotechnology
Delft University of Technology
2629 HZ Delft (The Netherlands)
[c] *Dr. T. Hayashi*
Science & Innovation Center
Mitsubishi Chemical Corporation
Aoba ward, Yokohama
Kanagawa 227-8502 (Japan)
[d] *M. Ligibel, Dr. E. Sager, Dr. K. Schroer, Dr. R. Snajdrova*
Novartis Institutes for BioMedical Research
Global Discovery Chemistry
4056 Basel (Switzerland)

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cctc.202201115

This publication is part of the Catalysis Talents Special Collection. Please check the ChemCatChem "homepage for more articles in the collection."

products, opening synthetic doors in medicinal chemistry research.

Going beyond the existing studies, we report herein that through enzyme engineering we could endow WelO5* variants with the ability to selectively chlorinate stereo-complementary martinelline-derived fragments directly from a synthesis mixture. Elucidation of the protein structure of WelO5* and one of its variants allowed us to obtain first insights into the factors determining substrate selectivity.

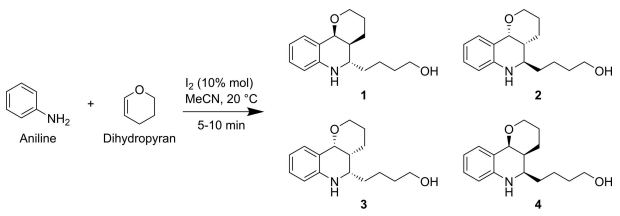
Results and Discussion

WelO5* variants enable racemic resolution of martinelline-derived *exo*-compounds

Substrate **1**, which is structurally related to the bradykinin receptor agonist martinelline, has been reported to possess anti-cancer activity.[10] Further functionalization of its molecular scaffold is thus desired for the exploration of the molecule's properties in SAR studies in medicinal chemistry research. With this in mind, Hayashi *et al.* evolved the αKGH WelO5* for the selective chlorination of **1**, leading to the enzyme variants CA2 (V81L/I161M) and CB2 able to produce the regio-divergent halogenated products **1a** and **1b**, respectively, when using purified substrate **1** in the biocatalysis mixture (Scheme S2).[9a]

The chemical synthesis route to the martinelline-derived fragment **1**, however, yields not only the desired molecule but leads to a mixture of *exo*- and *endo*-compounds (Scheme 1).[11] These compounds are present as pairs of racemates, requiring a preparative chiral separation for the isolation of the enantiomers.

We thus rationalized that to shorten the synthesis routes, it would be worthwhile to develop a late-stage functionalization scheme in which we could avoid the time-consuming purification step. Instead of separating the racemate pair of *exo*-compounds **1** and **2**, we aimed to exploit enzymatic substrate selectivity and use the *exo*-compound mixture directly for the biocatalysis reaction. Testing this setup, we were pleased to find that CB2 enabled the racemic resolution of an equimolar mixture of the *exo*-enantiomers **1** and **2** by exclusively chlorinating **1** to yield the target product **1b** (Figure 1, A1).



Scheme 1. Synthesis scheme of the four martinelline-derived fragments (**1**, **2**, **3**, and **4**), with the synthesis adapted from Lin *et al.*[11] The synthesis route yields in the racemate pairs of the *exo*- (**1** and **2**) and *endo*-compounds (**3** and **4**). With an *endo*/*exo* ratio of 34:66, the *exo*-compounds are the main products of the reaction.[11]

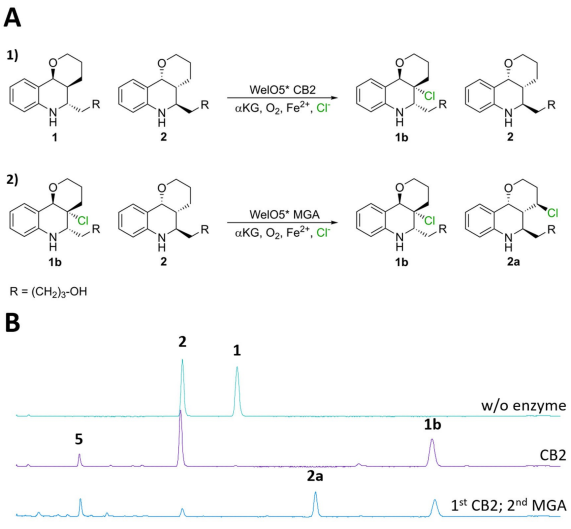


Figure 1. A) Reaction scheme of the selective chlorination of 1) **1** to **1b** by the CB2 variant and the subsequent reaction 2) **2** to **2a** by the MGA variant. For clarity, only substrate and chlorination product are depicted. B) LC-MS chromatogram (extracted ion count: *m/z* 250–350, positive mode) of the control without enzyme (green), the sequential biocatalysis with first CB2 (violet) and secondly MGA (blue). The reactions were performed with 40 μM purified CB2, 100 μM purified MGA, 0.25 mM of substrate **1** and **2** (0.5 mM in total), 106 mM α-ketoglutarate, 136 mM ascorbate, 550 mM NaCl, 1.7 mM ammonium iron(II) sulfate in 50 mM sodium phosphate buffer pH 8.

To enable full valorization of the two *exo*-enantiomers, we next decided to engineer a WelO5* variant capable to derivatize compound **2**. In contrast to the reported engineering campaign targeting compound **1**,[9a] we found that wildtype WelO5* did not exhibit any initial chlorination activity towards the stereoisomer **2**. Thus, with the aim to identify a halogenase which would accept the desired *exo*-compound **2**, we opted to screen enzyme libraries created in the WelO5* engineering campaign by Hayashi *et al.*[9a] Gratifyingly, this screening revealed the WelO5* double mutant LAV (V81L/I161V) to possess low chlorination activity towards **2**. Building on this modest biocatalyst, we next focused our engineering efforts on WelO5*'s known hotspots, namely sites V81, A88, and I161.[6c,9] To analyze these key residues combinatorially, we profited from a previously constructed enzyme library which was created by high-fidelity on-chip solid-phase gene synthesis limiting library diversity to the theoretical 8,000 variants. Screening 800 unique variants as confirmed by Sanger sequencing towards the chlorination of substrate **2** led to the identification of WelO5* triple mutant MGA (V81M/A88G/I161A) which, under screening conditions, yielded halogenation products **2a** (37%) and **2b** (<1%) as well as a hydroxylation product **2c** (5%). Remarkably, the inversion of substrate stereopreference compared to the WelO5* CB2 (V81R/I161S) variant was induced by the exchange of only three strategically placed amino acid positions in the active site. After performing the corresponding preparative scale biocatalysis reaction with MGA, we were able to elucidate the structure of the main chlorination product **2a** and the hydroxylation product **2c** by nuclear magnetic resonance (NMR, Table S4). The yield of the side-product **2b** was too small to

enable structural analysis, even after catalyzing the reaction with WelO5* WCP (V81W/A88C/I161P), a variant found to exhibit higher preference to produce this alternative product (Figure S1). Interestingly, NMR analysis showed that halogenase variant MGA functionalized the C11 atom of **2**, complementing the previously described derivatization sites for the *exo*-compounds, namely position C9 (variant CB2, product **1 b**) and position C12 (variant CA2, product **1 a**)[9a] (Scheme S1 and Table S4).

With variant MGA in hand, we followed up by establishing a sequential biocatalysis reaction enabling chlorination of **1** and **2** by first applying CB2 to yield **1 b** (94%) before adding MGA for the chlorination of **2** to **2 a** (87%; Figure 1B). In this way, the selective conversion of the main compounds from the synthesis, namely the *exo*-compounds **1** and **2**, could be converted into **1 b** and **2 a**, respectively. In this setup, we found that MGA could also accept **1 b**, the product of CB2's reaction, as denoted by its loss in the reaction mixture. In addition, a side product (**5**) with a *m/z* of 258 was identified in the LC-MS analysis and through NMR structure elucidation could be assigned to the aromatized compound formed from the martinelline-derived fragments (Table S4), potentially through enzymatic desaturation as previously described for other Fe(II)/αKG-dependent dioxygenases.[12] In this context, we can exclude a desaturation mechanism via elimination of the chloride installed at position C9 of product **1 b** as incubation of this compound with WelO5* variant CB2 did not yield the desaturation product **5** (Figure S6).

Selective chlorination of 1 directly from the synthesis mixture

Having established a first proof-of-concept with the *exo*-compounds, we decided to test the substrate selectivity of our enzyme variant even more stringently by attempting to chlorinate **1** to **1 b** directly from the synthesis mixture in which all four isomers (**1**–**4**) were present. Toward this goal, the four isomers were mixed in stoichiometric concentrations and applied as substrates for biocatalysis reactions with purified CB2 enzyme (Figure 2). The selective chlorination of **1** from this complex matrix (**1**–**4**) was achieved, yielding 90% of **1 b**. Again, we observed formation of the aromatized side product **5**, while at the same time a slight depletion of compounds **3** and **4** occurred indicating potential desaturation reactions.

Intrigued by the capability of WelO5* variants to distinguish between the stereoisomers **1** and **2**, we set out to solve the structure of WelO5* via protein crystallization. While crystal structures of the close enzyme homolog WelO5 have been solved for the substrate-free and -bound forms (PDB ID: 5IQS and 5IQT),[13] WelO5*, which stems from a different isolate from *Hapalosiphon welwitschii*, has not yet been successfully analyzed in this way. Notably, the here-investigated WelO5* naturally differs by 15 amino acids from WelO5, 11 of which can be found in a flexible C-terminal helix implicated in substrate recognition.[6a,b] Using Zn^{2+}, Cl^-, and α-ketoglutarate as additives, we succeeded to aerobically crystallize WelO5* in the *C*121 space group containing two protein chains in the asymmetric unit. The X-ray structure was solved at 2.3 Å resolution (PDB ID: 8ACV) using PDB ID: 5IQS (chain A) in molecular replacement experiments to solve the phases. In line with expectations, the structures of the two WelO5* chains resemble the eight-stranded β-sandwich fold as reported by Mitchell *et al.* for WelO5.[13] Interestingly, however, we found that by comparing our unbound WelO5* structure with the available substrate containing as well as free WelO5 complexes (PDB ID: 5IQT and 5IQS, respectively),[13] the elucidated WelO5* chains (A and B) adopted the closed-helix conformation which in WelO5 is only observed when the substrate is present in the active site with high occupation (Figure 3). In addition, WelO5*'s C-terminal helical region stands out due to the higher-than-

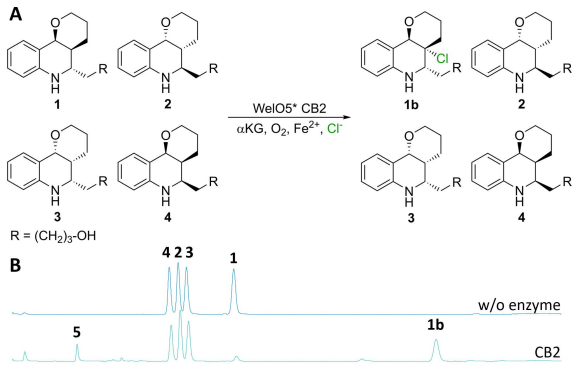


Figure 2. A) Reaction scheme of the selective chlorination **1** to **1 b** by the CB2 variant in the presence of the *exo*- and *endo*-compound mixture. For clarity, the substrate and chlorination product are depicted only. B) LC-MS chromatogram (total ion count: *m/z* 250–350, positive mode) of the control without enzyme (blue) and the CB2 reaction (green). The reaction was performed with 40 μM purified CB2, 0.25 mM substrate (1 mM in total), 106 mM α-ketoglutarate, 136 mM ascorbate, 550 mM NaCl, 1.7 mM ammonium iron(II) sulfate in 50 mM sodium phosphate buffer pH 8.

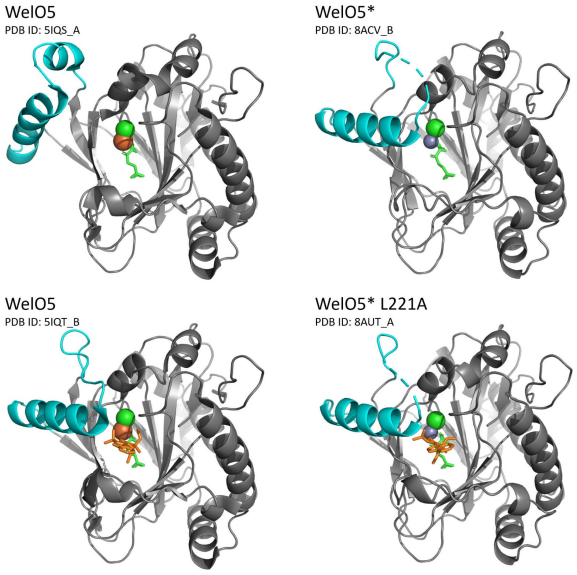


Figure 3. Comparison of the flexible C-terminal helices of the WelO5 (left) and WelO5* (right) structures without (top) and with (bottom) their native substrates. The protein is visualized as cartoon (grey) and the helix area (W210–Q238) is highlighted (cyan).

average B-factors (Figure S5) and a disordered section, supporting the flexible nature of this region.

As we exclusively observed the closed-helix conformation in our WelO5* structure, we decided to probe the effects of two additional residues on catalysis: L221 and I225, both located in the flexible helix, point directly into the active site when the enzyme is folded in this more compact structure. Using site-saturation mutagenesis in the context of the wildtype enzyme and MGA variant, we created the respective enzyme libraries using the degenerate NNK codon. Screening of the resulting enzyme libraries did not yield improved variants for the chlorination of **2** in the context of the MGA scaffold. However, introducing mutation L221A into the wildtype scaffold WelO5* led to an improved chlorination vs. hydroxylation ratio when converting the native substrate 12-*epi*-hapalindole C (**6**), although the overall yield of the chlorinated product was lower than for the wildtype (Figure S4).

Considering these results, we were interested to obtain substrate bound WelO5* structures to help us elucidate factors governing substrate recognition. Toward this end, we set up additional crystallization trials with WelO5* variants CB2, MGA, and L221A including substrates **1**, **2**, and **6**, respectively. We observed that all variants showed different crystallization behavior, with variant CB2 being the most difficult to crystallize. Unfortunately, all obtained CB2 and MGA crystals diffracted only poorly and the corresponding structures could not be solved. In contrast, the structure of the helix variant WelO5* L221A in complex with its native substrate **6** crystallized in the $P6_122$ space group with four chains in the asymmetric unit and was solved at 2.69 Å resolution (PDB ID: 8AUT). Here, electron density for the co-crystallized native substrate 12-*epi*-hapalindole C (**6**) was found in the active site of all four chains and the positioning of the substrate was found to be in good alignment with the placement of 12-*epi*-fischerindole U in the enzyme homolog WelO5 (PDB ID: 5IQT; Figure S3). In analogy to the WelO5* structure without substrate (*vide supra*), we found that the C-terminal helix region adopted the closed conformation in all four chains (Figure 3). In comparison to the WelO5* structure without substrate, structural rearrangement of the active site residues was only observed for I84, all other active site residues were found to retain wildtype geometry. Overall, the coordination of **6** was found to be facilitated via hydrophobic interactions by the surrounding side chains of the residues F77, V90, R153, H164, F169, I225, and F276, as visualized using the Protein-Ligand Interaction Profiler (PLIP) web tool[14] (Figure S3).

Enabled by the solved WelO5* crystal structure, we set out to further explore the factors governing the inversion of substrate stereopreference in WelO5* variants. To this end, we decided to model variants CB2 (V81R/I161S) and MGA (V81M/A88G/I161A) using UCSF Chimera,[15] and then dock the relevant substrates **1** and **2** via AutoDock Vina[16] into the respective active site (Figure 4). The substrate-protein interactions of the docked models were analyzed using the PLIP web tool[14] (Figure 4). This analysis revealed that both WelO5* variants, CB2 and MGA, bind ligands **1** and **2** via hydrophobic interactions, as expected due to the scarcity of heteroatoms in the structures of both substrates (Figure 4A, B). Further analyzing the docked

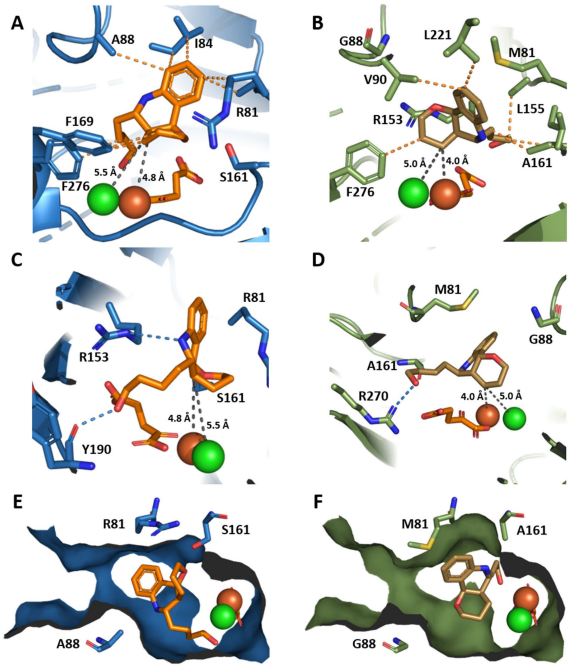


Figure 4. Docking poses of **1** and **2** in the active sites models of the CB2 (left, blue backbone) and MGA variants (right, green backbone), respectively. The hydrophobic interactions are annotated as dashed yellow lines (A and B), hydrogen bonds as dashed blue lines (C and D) and the distances to the key carbon atom and the Fe and Cl are shown as dashed grey lines. In addition, the surface of the active site is shown (bottom, E and F) with the mutations sites, substrates, and co-factors superimposed highlighting the differing active site shapes of the two variants.

model of CB2, two hydrogen bonds between heteroatoms in substrate **1** and amino acids in the CB2 active site were observed, namely between the guanidium moiety of R153 and the substrate nitrogen and between the terminal hydroxyl function of the substrate and the main chain carbonyl of Y190 (Figure 4C). Based on further visual inspection of our enzyme model a potential hydrogen bond between the hydroxyl function of the engineered serine (S161) with the ether oxygen atom of the substrate's tetrahydropyran ring was identified, which is presumed to additionally contribute to the positioning of the substrate. This hypothesis is further strengthened by the identical regioselectivity of WelO5* single mutants I161S and I161T underscoring the importance for a hydroxyl group at this position to support substrate placement within CB2 (V81R/I161S).[9a]

The second mutation in variant CB2, namely V81R, is likely to contribute to the coordination of the substrate by a cation-π interaction between the guanidine moiety of arginine and the phenyl function of the substrate, further contributing to the enzyme's specificity (Figure 4A).

Turning to the docked model of enzyme WelO5* MGA, which exhibits inverted substrate stereopreference, we identified a docking pose of substrate **2** which supports the regiospecific chlorination to **2a**. Besides the hydrophobic interaction (Figure 4B), a hydrogen bond between R270 and the terminal hydroxyl function was found. The newly introduced

alanine at position 161 participates in the hydrophobic interaction network to the substrate and additionally creates space in the active site for the accommodation of the substrate's phenyl function. Furthermore, replacement of valine with methionine at position 81 enables a potential coordination of the phenyl function via a methionine-π interaction. The subtle exchange of alanine to glycine at position 88 possibly also frees up additional space in the active site, although the mutation might not be needed for the mediation of the substrate acceptance. As the MGA variant was identified in a combinatorial library, information about the influence of single mutations is not available and the observed substrate specificity is conceivably caused by synergistic rather than linearly additive effects. Overall, our docked enzyme models reveal that the substantial re-shaping of the WelO5* active sites is a further decisive element to engineer variants with complementary substrate stereo-preference (Figure 4E, F).

Conclusion

In the quest to expand the biocatalytic toolbox for the halogenation of small molecules, the ability to modulate substrate specificity of Fe(II)/α-ketoglutarate-dependent halogenases is of key importance. Here, we extended the previous enzyme engineering studies of WelO5*[9a] by evolving the enzyme towards the halogenation of **2**, the enantiomer of an accepted martinelline-derived fragment. The modulation of the substrate selectivity was achieved by varying three key residues in the active site (V81, A88, and I161) yielding variant MGA characterized by the desired inverted substrate stereopreference. Applying the stereo-complementary halogenases CB2 and MGA enabled the selective chlorination of **1** and **2**, respectively, and allowed the racemic resolution of the enantiomers from their synthesis mixture. Structural data obtained from the protein crystal structures of WelO5* and substrate-bound WelO5* L221A variant provide insights into the factors governing substrate acceptance and underline the flexible nature of the C-terminal helix region which differentiates WelO5 homologs.

Overall, we show that freestanding halogenases can be encoded to exhibit exquisite substrate selectivity and regioselectivity affording a palette of derivatized products. In the quest to develop this valuable enzyme family as potent tools for asymmetric (late-stage) halogenation, an understanding of the decisive elements in non-natural substrate recognition is crucial and, as more such studies become available, will drive future engineering efforts.

Experimental Section

Plasmid constructs

The WelO5* gene with a terminal stop codon is subcloned in a pET28b vector between the NdeI and XhoI restriction sites, encoding the enzyme with an N-terminal His-tag. In addition, the WelO5* gene with a terminal stop codon was subcloned in a modified pET28b vector (pET28(mod)), where the nucleotide sequence between the NcoI and NdeI restrictions sites were removed and the NcoI was replaced by the NdeI restriction site, encoding WelO5* without His-tag. The nucleotide sequences of the open reading frames are given in the supporting information.

WelO5* library screening towards the chlorination of 2

The creation of the site-saturation library of the WelO5* gene in the pET28(mod) vector at positions V81, A88, and I161 is described elsewhere.[9b] The library was screened towards the chlorination of **2** in 96-deep-well-plate format. Preculture plates were prepared by filling 96-deep-well-plates with 1 mL well^{-1} LB-medium supplemented with 50 μg mL^{-1} kanamycin and inoculating the cultures with glycerol stocks of the library with a 96-well replicator. The preculture plates were incubated overnight at 37 °C, 300 rpm (Duetz-System in Kuhner shaker, 50 mm shaking diameter). For the inoculation of the expression plates, 100 μL well^{-1} of the precultures were transferred into 900 μL well^{-1} ZYM-5052 autoinduction media supplemented with 50 μg mL^{-1} kanamycin and incubated overnight at 20 °C, 300 rpm (Duetz-System in Kuhner shaker, 50 mm shaking diameter). The expression plates were harvested as cell pellets by centrifugation for 20 min, 3,428 *g*, 4 °C, the supernatant was discarded, and the plates with the cell pellets were stored at −20 °C. For the activity screening, the cell pellets were resuspended and lysed in 100 μL well^{-1} lysis buffer comprising 1 mg mL^{-1} lysozyme, 0.5 mg mL^{-1} polymyxin B sulfate, and DNase I in 50 mM sodium phosphate buffer pH 8 and incubated for 20 min, 20 °C, 1200 rpm (ThermoMixer C, Eppendorf, Germany). Using previously optimized reaction conditions,[9] the biocatalysis reaction was initiated by the addition of 100 μL well^{-1} assay mix comprising 1 mM **2**, 212 mM α-ketoglutarate disodium dihydrate, 272 mM (+)-sodium L-ascorbate, 1100 mM NaCl, 3.4 mM ammonium iron(II) sulfate hexahydrate in 50 mM sodium phosphate buffer pH 8 and incubated for 90 min, 800 rpm, 20 °C (ThermoMixer C, Eppendorf, Germany). The reaction was quenched by the addition of 800 μL 62.5 % methanol and incubated for 15 min, 800 rpm, 20 °C and subsequently centrifuged for 20 min, 3,428 *g*, 4 °C. The reaction was analyzed by injecting 2 μL of the supernatant in the LC-MS (Agilent 1290 InfinityLab LC/MSD; Agilent, USA) equipped with the InfinityLab Poroshell 120 EC–C18 column (3.0×50 mm, 2.7 μm, Agilent, USA). The mobile phases A and B were ddH_2O + 5 % acetonitrile + 0.2 % formic acid and acetonitrile + 0.2 % formic acid, respectively. The following method was used at 40 °C and a flow rate of 0.6 mL min^{-1}: 10 % B for 1 min, gradient from 10 to 50 % B in 0.5 min, 50 % B for 2 min, gradient from 50 to 85 % B in 0.5 min, 85 % B for 1 min. The substrate and products were detected in the DAD (254 nm), MSD scan (*m/z* 150–600), and MSD SIM (*m/z* $[M+H]^+$) at the following retention times: **2** at 2.29 min, **2a** at 2.75 min, **2b** at 2.55 min, **2c** at 2.00 min.

Preparative-scale biocatalysis of 2 and product structure elucidation

The preparative biocatalysis was performed by scaling up the above-described screening reaction. For the expression, a preculture with 150 mL LB-medium supplemented with 50 μg mL^{-1} kanamycin was inoculated with the respective glycerol stock of the WelO5* variant and incubated overnight at 37 °C, 160 rpm. For the inoculation of the expression culture, 100 mL of preculture was transferred in 900 mL ZYM-5052 autoinduction medium in two 2 L baffled flasks and incubated overnight at 20 °C, 140 rpm (50 mm shaking diameter). The cells were harvested by centrifugation for 20 min at 3,428 g, 4 °C, and the cell pellet was stored at −20 °C. The cells were resuspended and lysed by the addition of 40 mL lysis

buffer (50 mM sodium phosphate buffer pH 8, 1 mg mL^{-1} lysozyme, 0.5 mg mL^{-1} polymyxin B-sulfate, and DNase I) in a 500 mL baffled flask and incubated at 20 °C, 120 rpm for 20 min. The biocatalytic reaction was initiated by the addition of 40 mL assay mix (2 mM **2** (20.9 mg **2** in 1 mL methanol), 220 mM α-ketoglutarate disodium salt dihydrate, 212 mM (+)-sodium L-ascorbate, 1000 mM NaCl, 2.6 mM ammonium iron(II) sulfate hexahydrate in 50 mM sodium phosphate buffer pH 8) and incubated for 16 h at 20 °C, 120 rpm. The reaction was quenched with 80 mL methanol and clarified by centrifugation for 40 min at 3,428 g, 4 °C. The methanol was removed from the clarified supernatant with a rotary evaporator (Büchi, Switzerland) and the substrates and products were extracted two times by the addition of 200 mL ethyl acetate each. The organic phase was washed with 400 mL brine solution and dried with Na_2SO_4. The ethyl acetate was removed from the washed and dried extract with a rotary evaporator and a brown, viscous oil was obtained. Crude reaction mixture was dissolved in 2–3 mL DMSO and injected on the preparative HPLC, using Nucleodur C18 column. The mobile phases A (ddH_2O + 1% formic acid) and B (acetonitrile + 1% formic acid) were applied with a flow rate of 40 mL min^{-1} and mixed with following regime: 0% B for 5 min; 0 to 65% B gradient over 30 min, 65% of B for 5 min. The substrate and products were detected in the DAD (200–600 nm), with peak selection at 300 nm. Fractions containing each product were collected, lyophilized, and submitted for NMR analysis (see SI).

Creation, expression, and purification of WelO5* His-tag variants

Since the screening of the WelO5* variants was performed in the pET28(mod) vector in absence of a His-tag, the recloning of beneficial variants in a pET28b vector with an N-terminal His-tag was required for His-tag purification. Therefore, the gene was amplified with flanking primers (fw: 5′-GTG AGC GGA TAA CAA TTC CCC TCT AG-3′; rv: 5′-GCT TTG TTA GCA GCC GGA TCT CAG-3′), digested with NdeI and XhoI, and ligated in a pET28b vector digested with the same enzymes. The recloning was confirmed by Sanger sequencing (Microsynth AG, Switzerland). Alternatively, the recloning in pET28b was performed with MegaWhop PCR with the His-tagged WelO5* gene as a template.[17] The N-terminal His-tagged variants were transformed in *E. coli* BL21(DE3) and glycerol stocks were stored at −80 °C. For the expression of the His-tagged WelO5* variant, 150 mL LB-medium supplemented with 50 μg mL^{-1} kanamycin were inoculated with the corresponding glycerol stock as preculture and incubated overnight at 37 °C, 160 rpm. For the inoculation of the expression culture, 100 mL of the preculture was transferred into 900 mL ZYM-5052 autoinduction media supplemented with 50 μg mL^{-1} kanamycin in two baffled 2 L flasks and incubated overnight at 20 °C, 140 rpm (50 mm shaking diameter). The culture was harvested as a cell pellet by centrifugation for 20 min, 3,428 *g*, 4 °C, the supernatant was discarded and the pellet was stored at -20 °C. For purification, the cell pellet was resuspended in 20 mL lysis buffer (10 mM imidazole, 10 mM β-mercaptoethanol, 500 mM NaCl, 0.1% Tween-20, 50 mM TRIS–HCl pH 7.4 and DNase I), lysed by ultrasonic treatment (2x 1 min with 2/sec intervals on ice, Sonopuls, Bandelin, Germany), and clarified by centrifugation for 1 h at 20,133 *g*, 4 °C. The immobilized metal affinity chromatography was performed by loading the clarified lysate on a pre-equilibrated column with 5 mL Ni–NTA His-beads (FastSep NTA Agarose, MCLAB, USA), washing the beads with 60 mL washing buffer (20 mM imidazole, 10 mM β-mercaptoethanol, 500 mM NaCl, 50 mM TRIS–HCl pH 7.4), and eluting the enzyme in 1 mL fractions with 15 mL elution buffer (200 mM imidazole, 10 mM β-mercaptoethanol, 500 mM NaCl, 50 mM TRIS–HCl pH 7.4). The protein-containing fractions were combined and concentrated (Amicon Ultra 10,000 NMWL, Merck Millipore Ltd., Ireland). Depending on the required purity of the WelO5* variants, size-exclusion chromatography (SEC) was performed with an ÄKTA pure system (GE Healthcare Life Sciences, USA) equipped with the HiLoad 16/600 Superdex 75 pg column (GE Healthcare Life Sciences, USA) in SEC-buffer (200 mM NaCl, 50 mM TRIS–HCl pH 7.4). The WelO5* variants were desalted in 50 mM sodium phosphate buffer pH 8 using PD-10 desalting columns (GE Healthcare, United Kingdom) according to the manufacturer's protocol, concentrated, shock-frozen with liquid nitrogen, and stored at −80 °C.

Biocatalysis reaction, workup and chiral LC-MS analysis

The biocatalysis reactions were performed with purified WelO5* variants (40 and 100 μM for CB2 and MGA, respectively), 0.25 mM of each substrate (**1**–**4**), 106 mM α-ketoglutarate disodium dihydrate, 136 mM (+)-sodium L-ascorbate, 550 mM NaCl, 1.7 mM ammonium iron(II) sulfate hexahydrate in 50 mM sodium phosphate buffer pH 8 in 200 μL format.

The biocatalysis reactions were quenched by the addition of 800 μL 62.5% methanol, incubated for 20 min at 1300 rpm (ThermoMixer F0.5, Eppendorf), and centrifuged at 21,700 *g* for 3 min. The supernatant was transferred to a glass vial for analysis. The chiral LC-MS analysis was performed by injecting 2 μL in the Agilent 1260 HPLC system with a single quadrupole MSD (Agilent, USA) and the Chiralcel OD–RH (150 mm × 4.6 mm, 5.0 μm; Daicel Corporation, Japan) column at 30 °C. The mobile phases A and B were ddH_2O + 5% acetonitrile + 0.2% formic acid and acetonitrile + 0.2% formic acid, respectively. The following method was used at a flow rate of 0.5 mL min^{-1}: 10% B for 1 min, gradient from 10 to 30% B in 6 min, gradient from 30 to 60% B in 25 min, gradient from 60 to 90% B in 3 min, and 90% B for 5 min. The substrates and products were detected in the DAD (254 nm), MSD scan (*m/z* 250–350), and MSD SIM (*m/z* $[M+H]^+$) at the following retention times: **1** at 18.57 min, **1b** at 31.50 min, **2** at 14.97 min, **2a** at 23.80 min, **3** at 15.53 min, **4** at 14.37 min, **5** at 8.27 min.

Protein crystallization and structure elucidation

The purified WelO5* variants containing an N-terminal His_6-tag were diluted to 20 mg mL^{-1} in crystallization buffer (0.58 mM $ZnCl_2$, 2.9 mM α-ketoglutarate disodium salt dihydrate). In the samples prepared for co-crystallization, 20 mM **1**, **2**, or **6** was added from a 400 mM stock in DMSO. The samples were mixed with crystallization reagent (22.14% PEG-4000, 30 mM sodium acetate pH 4.5 and 70 mM Bis-TRIS pH 5.5 or 30% PEG-4000, 200 mM $LiSO_4$, in 100 mM TRIS pH 8.5 for the wildtype or L221A variant, respectively) and sitting drop vapor diffusion crystallization was performed aerobically in Intelli-Plate 96-3 LVR plates (Hampton Research, USA). The reported crystallization conditions for the variants were identified after the investigation of several screening plates. The plates were incubated in the Rock Imager 1000 (Formulatrix, USA) and monitored via the Rock Maker Software (Formulatrix, USA) at the Protein Crystallization Center (University of Zurich, Switzerland). Needle- and rhomboid-shaped crystals were obtained for WelO5* and the L221A variant, respectively, after incubation at 20 °C for 10 days (Figure S2). The crystals were collected by overlaying the crystallization well with 30% of glycerol (in reagent buffer from the reservoir) as a cryoprotectant, picking the crystals with a LithoLoop (Molecular Dimensions, United Kingdom), and shock-freezing the sample in liquid nitrogen. The crystals were measured at the PXI or PXIII beamline at the SLS synchrotron (Paul Scherrer Institute, Switzerland). The diffraction data were integrated with XDS[18] and solved by molecular replacement using MOLREP[19] with the WelO5 crystal structure PDB ID: 5IQS (chain A)[13] for the WelO5* wildtype structure. The subsequent WelO5* L221A structure was solved with

the obtained WelO5* wildtype crystal structure. Iterative refinement and model building steps were performed with BUSTER[20] or Refmac[21] (for WelO5* wildtype and WelO5* L221A, respectively) and COOT.[22] The programs (with exception of BUSTER) were embedded in the CCP4i2 suite.[23]

In silico variant generation and molecular docking

The WelO5* variant structures for molecular docking were generated *in silico* based on the obtained WelO5* wildtype crystal structure. The WelO5* chain A structure was imported in UCSF Chimera 1.13.1,[15] the Zn^{2+} ion was replaced by Fe^{2+}, the respective amino acids mutated, and energy minimization performed to yield the variant structure. The molecular docking was performed with AutoDock Vina 1.2.2[16] with the ligand created in Chem3D 16.0 (PerkinElmer, USA) and the prepared protein structure as the receptor. The docking was performed with the Vina scoring function and the following custom settings: output modes 50, exhaustiveness 100 with residues surrounding the active site set as flexible. The docking results were inspected, and the images created in PyMOL 2.5.3 (Schrödinger LLC., USA). The docking poses were evaluated based on the positioning of the ligand in the active site to explain the observed product formation. The binding affinities for the selected poses of CB2 with **1** and MGA with **2** were $-20.8\ kcal\,mol^{-1}$ ($-25.0\ kcal\,mol^{-1}$ for highest ranked pose) and $-23.6\ kcal\,mol^{-1}$ ($-24.1\ kcal\,mol^{-1}$ for highest ranked pose), respectively. The docked models were evaluated using the PLIP Protein-Ligand Interaction Profiler (PLIP) web tool[14] in addition to manual inspection. The protein models with the docked substrates, output and log files from the docking were deposited at the open-access online repository Zenodo (zenodo.org) and the following DOI was assigned: 10.5281/zenodo.7157068.

Substrate synthesis

The synthesis of the martinelline-derived fragments was performed as described by Hayashi *et al.*[9a] analog to the protocol by Lin *et al.*[11] The synthesis of 12-*epi*-hapalindole C (**6**) was performed at Novartis' external partner by custom synthesis and the NMR is included in SI.

Acknowledgements

This work was created as part of NCCR Catalysis, a National Centre of Competence in Research funded by the Swiss National Science Foundation (Grant number 180544). We are grateful for the support in protein crystallization by the Protein Crystallization Center of Zurich University and the Paul Scherrer Institute for measuring time at the Swiss light source (SLS) synchrotron.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

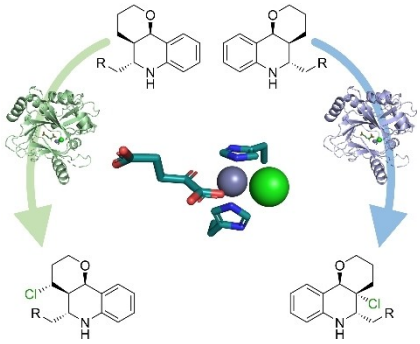
Keywords: biocatalysis · halogenase · protein engineering · kinetic resolution · protein crystallography

[1] a) J. Büchler, A. Papadopoulou, R. Buller, *Catalysts* **2019**, *9*, 1030; b) J. Latham, E. Brandenburger, S. A. Shepherd, B. R. K. Menon, J. Micklefield, *Chem. Rev.* **2018**, *118*, 232–269.
[2] L. Guillemard, N. Kaplaneris, L. Ackermann, M. J. Johansson, *Nat. Chem. Rev.* **2021**, *5*, 522–545.
[3] A. E. Fraley, D. H. Sherman, *Bioorg. Med. Chem. Lett.* **2018**, *28*, 1992–1999.
[4] a) J. T. Payne, M. C. Andorfer, J. C. Lewis, in *Methods in Enzymology, Vol. 575* (Ed.: S. E. O'Connor), Academic Press, **2016**, pp. 93–126; b) H. Minges, N. Sewald, *ChemCatChem* **2020**, *12*, 4450–4470.
[5] M. L. Hillwig, X. Liu, *Nat. Chem. Biol.* **2014**, *10*, 921–923.
[6] a) Q. Zhu, X. Liu, *Beilstein J. Org. Chem.* **2017**, *13*, 1168–1173; b) M. L. Hillwig, Q. Zhu, K. Ittiamornkul, X. Liu, *Angew. Chem. Int. Ed.* **2016**, *55*, 5780–5784; *Angew. Chem.* **2016**, *128*, 5874–5878; c) S. Duewel, L. Schmermund, T. Faber, K. Harms, V. Srinivasan, E. Meggers, S. Höbenreich, *ACS Catal.* **2020**, *10*, 1272–1277.
[7] M. E. Neugebauer, K. H. Sumida, J. G. Pelton, J. L. McMurry, J. A. Marchand, M. C. Y. Chang, *Nat. Chem. Biol.* **2019**, *15*, 1009–1016.
[8] C. Zhao, S. Yan, Q. Li, H. Zhu, Z. Zhong, Y. Ye, Z. Deng, Y. Zhang, *Angew. Chem. Int. Ed.* **2020**, *59*, 9478–9484.
[9] a) T. Hayashi, M. Ligibel, E. Sager, M. Voss, J. Hunziker, K. Schroer, R. Snajdrova, R. Buller, *Angew. Chem. Int. Ed.* **2019**, *58*, 18535–18539; *Angew. Chem.* **2019**, *131*, 18706–18711; b) J. Büchler, S. H. Malca, D. Patsch, M. Voss, N. J. Turner, U. T. Bornscheuer, O. Allemann, C. Le Chapelain, A. Lumbroso, O. Loiseleur, R. Buller, *Nat. Commun.* **2022**, *13*, 371.
[10] P. Y. Chung, J. C. O. Tang, C. H. Cheng, Z. X. Bian, W. Y. Wong, K. H. Lam, C. H. Chui, *Springerplus* **2016**, *5*, 271.
[11] X. F. Lin, S. L. Cui, Y. G. Wang, *Tetrahedron Lett.* **2006**, *47*, 4509–4512.
[12] a) F. Meyer, R. Frey, M. Ligibel, E. Sager, K. Schroer, R. Snajdrova, R. Buller, *ACS Catal.* **2021**, *11*, 6261–6269; b) A. Papadopoulou, F. Meyer, R. M. Buller, *Biochemistry* **2022**; c) N. P. Dunham, W. C. Chang, A. J. Mitchell, R. J. Martinie, B. Zhang, J. A. Bergman, L. J. Rajakovich, B. Wang, A. Silakov, C. Krebs, A. K. Boal, J. M. Bollinger, *J. Am. Chem. Soc.* **2018**, *140*, 7116–7126; d) T. Y. Chen, S. Xue, W. C. Tsai, T. C. Chien, Y. S. Guo, W. C. Chang, *ACS Catal.* **2021**, *11*, 278–282; e) H. J. Liao, J. K. Li, J. L. Huang, M. Davidson, I. Kurnikov, T. S. Lin, J. L. Lee, M. Kurnikova, Y. S. Guo, N. L. Chan, W. C. Chang, *Angew. Chem. Int. Ed.* **2018**, *57*, 1831–1835; *Angew. Chem.* **2018**, *130*, 1849–1853.
[13] A. J. Mitchell, Q. Zhu, A. O. Maggiolo, N. R. Ananth, M. L. Hillwig, X. Y. Liu, A. K. Boal, *Nat. Chem. Biol.* **2016**, *12*, 636–640.
[14] M. F. Adasme, K. L. Linnemann, S. N. Bolz, F. Kaiser, S. Salentin, V. J. Haupt, M. Schroeder, *Nucleic Acids Res.* **2021**, *49*, W530–W534.
[15] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605–1612.
[16] J. Eberhardt, D. Santos-Martins, A. F. Tillack, S. Forli, *J. Chem. Inf. Model.* **2021**, *61*, 3891–3898.
[17] K. Miyazaki, in *Methods in Enzymology, Vol. 498* (Ed.: C. Voigt), Academic Press, **2011**, pp. 399–406.
[18] W. Kabsch, *Acta Crystallogr. Sect. D* **2010**, *66*, 125–132.
[19] A. Vagin, A. Teplyakov, *J. Appl. Crystallogr.* **1997**, *30*, 1022–1025.
[20] G. Bricogne, E. Blanc, M. Brandl, C. Flensburg, P. Keller, W. Paciorek, P. Roversi, A. Sharff, O. S. Smart, C.Vonrhein, W. T. O., BUSTER version 2.10.3, United Kingdom: Global Phasing Ltd., **2017**.
[21] G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallogr. Sect. D* **1997**, *53*, 240–255.
[22] P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D* **2004**, *60*, 2126–2132.
[23] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, *Acta Crystallogr. Sect. D* **2011**, *67*, 235–242.

Manuscript received: September 13, 2022
Revised manuscript received: October 17, 2022
Accepted manuscript online: October 24, 2022
Version of record online: ■■, ■■■

RESEARCH ARTICLE

Enzymatic kinetic resolution by variants of the halogenase WelO5* allows for the selected chlorination of two martinelline-derived fragments. Strikingly, these engineered enzyme variants differ by only three strategically selected amino residues in the active site. In addition, the protein structure elucidation of WelO5* provides information about the substrate coordination of this enzyme and underlines the flexible nature of the C-terminal helix.



*Dr. M. Voss, S. Hüppi, D. Schaub, Dr. T. Hayashi, M. Ligibel, Dr. E. Sager, Dr. K. Schroer, Dr. R. Snajdrova, Prof. R. Buller**